Eur J Plant Pathol (2009) 124:45–55
DOI 10.1007/s10658-008-9390-2

Molecular characterisation of *Turnip mosaic virus* isolates from Brassicaceae weeds

Shirin Farzadfar · Yasuhiro Tomitaka · Mutsumi Ikematsu · Ali Reza Golnaraghi · Reza Pourrahim · Kazusato Ohshima

Received: 29 July 2008 / Accepted: 13 October 2008 / Published online: 30 October 2008

Abstract Eight provinces of Iran were surveyed during 2003–2008 to find Brassicaceae reservoir weed hosts of *Turnip mosaic virus* (TuMV). A total of 532 weed samples were collected from plants with virus-like symptoms. The samples were tested for the presence of TuMV by enzyme-linked immunosorbent assay using specific antibodies. Among those tested, 340 samples (64%) were found to be infected with TuMV. *Rapistrum rugosum*, *Sisymberium loeselii*, *S. irio* and *Hirschfeldia incana* were identified as the Brassicaceae weed hosts of TuMV, and the former two plant species were found to be the most important weed hosts for the virus in Iran. The full-length sequences of the genomic RNAs of IRN TRa6 and IRN SS5 isolates from *R. rugosum* and *S. loeselii* were determined. No evidence of recombination was found in both isolates using different recombination-detecting programmes. Phylogenetic analyses of the weed isolates with representative isolates from the world showed that the IRN TRa6 and IRN SS5 isolates fell into an ancestral basal-*Brassica* group. This study shows for the first time the wide distribution and phylogenetic relationships of TuMV from weeds in the mid-Eurasia of Iran.

Keywords *Turnip mosaic virus* · Brassicaceae · Weeds · Iran · Evolution · Phylogeny

S. Farzadfar (✉) · R. Pourrahim
Department of Plant Virology,
Plant Pests and Diseases Research Institute,
P.O. Box 19395-1454, Tehran, Iran
e-mail: farzadfar2002@yahoo.com

Y. Tomitaka · M. Ikematsu · K. Ohshima (✉)
Laboratory of Plant Virology, Faculty of Agriculture,
Saga University,
1-banchi, Honjo-manchi,
Saga 804-8502, Japan
e-mail: ohshima@cc.saga-u.ac.jp

A. R. Golnaraghi
Department of Plant Protection,
College of Agriculture and Natural Resources,
Science and Research Branch, Islamic Azad University,
P.O. Box 14515-775, Tehran, Iran

Introduction

Turnip mosaic virus (TuMV) belongs to the genus *Potyvirus* within the family of *Potyviridae* (Fauquet et al. 2005) and is probably the most common and important virus in Brassicaceae crops throughout the world. TuMV is widespread in both temperate and subtropical regions of Africa, Asia, Europe, Oceania and North and South America (Provvidenti 1996; Ohshima et al. 2002). TuMV is the second most important virus infecting field-grown vegetable crops after *Cucumber mosaic virus* (CMV) (Tomlinson 1987; Walsh and Jenner 2002). TuMV, like other potyviruses, is transmitted by aphids in the non-persistent manner, and has flexuous filamentous particles of 700–750 nm in length, each of which contains a single-stranded positive-sense RNA molecule of *ca.* 10,000 nucleo-

tides (Provvidenti 1996; Fauquet et al. 2005). Potyvirus genomes have a single open reading frame (ORF) flanked by terminal untranslated regions. The ORF is translated to a single large polyprotein, which is hydrolysed into at least ten proteins by virus-encoded proteinases (Urcuqui-Inchima et al. 2001).

Previous studies have shown that TuMV isolates had been biologically classified into four host-infecting types; host type [(B)]: isolates infected *Brassica* plants latently and occasionally and did not infect *Raphanus* plants, host type [B]: isolates infected many of *Brassica* plants systemically with mosaic on uninoculated leaves but did not infect *Raphanus* plants, host type [B(R)]: isolates infected many of *Brassica* systemically with mosaic on uninoculated leaves but infected *Raphanus* plants latently and occasionally, and host type [BR]: isolates infected both of *Brassica* and *Raphanus* plants systemically with mosaic on uninoculated leaves. Moreover, phylogenetic analysis using different virus isolates collected from around the world revealed four main TuMV genogroups called basal-*Brassica* (basal-B), basal-*Brassica*/*Raphanus* (basal-BR), Asian-*Brassica*/*Raphanus* (Asian-BR) and world-*Brassica* (world-B) (Ohshima et al. 2002; Tomimura et al. 2004). The basal-B cluster of (B) or B-host type isolates was most variable, was not monophyletic, and came from both non-brassicas and brassicas. The studies also have shown that the different TuMV subpopulations have probably emerged from the more ancient Eurasian subpopulations such as those found in southeast Europe, Asia Minor and mid-Eurasia regions (Ohshima et al. 2002; Tomimura et al. 2004). In these regions, *Brassica* crops are an important component of local agriculture; in Europe, the crops are mostly *Brassica* species; and in Asia Minor and mid-Eurasia regions, both *Brassica* and *Raphanus* species are important. Although TuMV was recently reported in Iran (Farzadfar et al. 2005), little data on the biological and molecular characteristics are available from the mid-Eurasia regions. Moreover, previous studies have reported the occurrence of TuMV on different species of cultivated Brassicaceae plants; therefore that of TuMV on Brassicaceae weeds is largely unknown to date, even though reservoir weed hosts are important for virus expansion (Ohshima et al. 2002, 2007).

In the present study, the occurrence of TuMV is reported from different Brassicaceae weed species. Furthermore, two full genomes of these isolates were sequenced and their phylogenetic relationships with worldwide isolates are discussed.

Materials and methods

Collection of weed samples

Weed samples, including *Brassica deflexa* (persian-kaali), *Hirschfeldia incana* (shortpod mustard) *Rapistrum rugosum* (rugose rapistrum), *Sinapis arvensis* (charlock), *Sisymberium irio* (London rocket), and *S. loeselii* (small tumbleweed-mustard), were collected from the Brassicaceae crop-producing areas of eight provinces of Iran during the growing seasons of 2003–2008 (Fig. 1, Table 1). Samples were collected from the weeds showing leaf deformation, mosaic, necrosis, rugosity, stunting, vein yellows and yellowing. The collected leaves were immediately placed in plastic bags and transported to the laboratory. They were kept at 4°C until use.

Serological virus identification

Collected samples were tested for TuMV by double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA), as described previously by Clark and Adams (1977), using specific polyclonal antibodies purchased from Loewe (Sauerlach, Germany). Briefly, microtitre plates (Maxisorb, NUNC, Denmark) were coated with 1:200 TuMV-IgG in carbonate buffer. Leaves were ground with a mortar and pestle in extraction buffer (1:5 wt/vol), and leaf extracts were then added to the plates in duplicate wells and incubated overnight at 4°C. The presence of TuMV in the samples was detected by TuMV-specific antibody conjugated to alkaline phosphatase using *p*-nitrophenyl phosphate substrate. Absorbance values at 405 nm were measured using a Multiscan-334 microtiter plate reader (Lab system, Finland). Absorbance values, measured 60 min after adding the substrate, greater than three times those of the negative controls were considered positive.

Virus isolates and host range

Weed samples collected from Iran were homogenised in 0.1 M Na-phosphate buffer (pH 7.4) containing 0.02% 2-mercaptoethanol and mechanically inoculat-

Fig. 1 Map of Iran showing the location of provinces that were surveyed during the crop growing seasons of 2003–2008

ed onto turnip (*Brassica rapa* cv. Hakatasuwari) or *Nicotiana benthamiana* dusted with carborundum. Altogether, 19 out of 340 ELISA-positive TuMV isolates were selected, considering differences for origin and host plants. Each virus isolate was serially cloned through single isolations at least three times using *Chenopodium amaranticolor* or *C. quinoa*, and they were subsequently propagated in turnip or *N. benthamiana*. First, all 19 isolates were inoculated onto *B. napus* cv. Otsubu, *B. pekinensis* cvs Kyoto-3go and Nozaki-1go, *B. rapa* cv. Hakatasuwari, and *Raphanus sativus* cvs Akimasari and Taibyo-sobutori, and the host-infecting types of each isolate were roughly evaluated. Because most isolates showed similar host reactions to the plant species we examined, two selected isolates, IRN TRa6 and IRN SS5 from *R. rugosum* and *S. loeselii* in Tehran and Semnan provinces, were used for the detailed host-range studies. The isolates were inoculated onto young plants of the families Asteraceae, Balsaminaceae, Brassicaceae, Chenopodiaceae and Solanaceae. In particular, Brassicaceae plants, for instance *B. oleracea*, *B. rapa*, *Eruca sativa* and *R. sativus*, of different seed origin from many countries were used. In addition, some Brassicaceae weeds including *Hirschfeldia incana*, *R. rugosum*, *Sinapis alba*, *S. arvensis*, *Sisymberium irio* and *S. loeselii* were also tested (Table 2). Inoculated plants were kept in a greenhouse with at least 16 h light, 25±5°C and 50–70% relative humidity. The symptoms were observed periodically for more than 1 month, after which the presence of virus in the inoculated plants was examined by DAS-ELISA.

Extraction of viral RNA and sequencing

Total RNA was extracted from TuMV-infected *B. rapa* or *N. benthamiana* using Isogen (Nippon Gene, Tokyo, Japan) and first-strand cDNA synthesis was performed using PrimeScript™ reverse transcriptase (Takara Bio Inc., Otsu, Japan), according to the manufacturers' instructions. Polymerase chain reaction (PCR) amplifications were done using high-fidelity Platinum™ *Pfx* DNA polymerase (Invitrogen,

Table 1 Occurrence of *Turnip mosaic virus* (TuMV) in Brassicaceae weed hosts in different provinces of Iran[a]

Province	Host (common name)	No. of samples infected/collected	Collection year
Fars	*Sisymberium loeselii* (Small tumbleweed-mustard)	0/17 (0.0)[b]	2003
Hamedan	Rapistrum rugosum (Rugose rapistrum)	13/13 (100)	2005
	Sisymberium loeselii	70/70 (100)	2005
	Sisymberium irio (London rocket)	9/22 (40.9)	2005, 2008
Kerman	Brassica deflexa (Persiankaali)	0/10 (0.0)	2005
	Sinapis arvensis (Charlock/wild mustard)	0/10 (0.0)	2006
	Sisymberium irio	0/30 (0.0)	2006
	Sisymberium loeselii	27/60 (45.0)	2005, 2006
Khorasan	Rapistrum rugosum (Rugose rapistrum)	5/5 (100)	2005
	Sisymberium loeselii	24/30 (80.0)	2004
Markazi	Sisymberium loeselii	35/35 (100)	2003, 2006
Mazandaran	Sinapis arvensis	0/20 (0.0)	2005
Semnan	*Sisymberium loeselii*	10/10 (100)	2003
Tehran	Hirschfeldia incana (Shortpod mustard)	1/5 (20.0)	2006
	Rapistrum rugosum	64/64 (100)	2004, 2005, 2006
	Sinapis arvensis	0/7 (0.0)	2004
	Sisymberium irio	13/24 (54.2)	2004, 2006
	Sisymberium loeselii	69/100 (69.0)	2004, 2005, 2006
Total[c]		340/532 (63.9)	

[a] identification is based on serological reactions (ELISA).

[b] percent of virus infection rate in the symptomatic weed samples collected.

[c] average of virus infection.

Carlsbad, CA, USA). PCR products were separated by electrophoresis in agarose gels. The expected fragments were excised from the gels, cleaned by the QIAquick Gel Extraction Kit (Qiagen K.K., Japan) and cloned into pBluescript II SK$^+$. Nucleotide sequences from each isolate were determined using overlapping cloned DNAs, and also several independent RT-PCR products to cover the full genome. The sequences of adjacent regions of the genome were overlapped by at least 100 bp to ensure that they were from the same genome and were not from different components of a genome mixture. DNA sequencing was done by primer walking in both directions using the BigDye Terminator v3.1 Cycle Sequencing Ready Reaction Kit and an Applied Biosystems Genetic Analyser DNA model 310 (Applied Biosystems, Foster City, CA, USA). Nucleotide sequence data were assembled using BioEdit Version 5.0.9 (Hall 1999).

Recombination analyses

The genomic sequences of the 99 isolates available from the international sequence databases and two Iranian isolates from weeds analysed in this study were used for evolutionary analyses. Two sequences of *Japanese yam mosaic virus* (JYMV) (GenBank Accession codes: AB027007 and AB016500), one of *Scallion mosaic virus* (ScMV) (AJ316084), and one of *Narcissus yellow stripe virus* (NYSV) (AM158908) were used for outgroups as BLAST searches identified those to be the sequences in the international sequence databases most closely and consistently related to those of TuMV. TuMV protein 1 (P1) genes were more closely related to those of JYMV than ScMV, whereas for some other genomic regions between the helper-component proteinase protein (HC-Pro) and nuclear inclusion b protein (NIb) sequences, the opposite was true, while the TuMV coat protein (CP) gene was most closely related to that of NYSV. All 99 TuMV P1 sequences were aligned with those of the two JYMV isolates as outgroups, the CP sequences with that of NYSV, and the remaining sequences with those of JYMV and ScMV using CLUSTAL X (Jeanmougin et al. 1998). However, this procedure resulted in some gaps that were not in multiples of three nucleotides. Therefore, the amino acid sequences corresponding to individual regions were aligned with the appropriate outgroups indicated above using CLUSTAL X with TRANS-

Table 2 Symptoms induced by two weed isolates of *Turnip mosaic virus* on indicator plants

Plant	Common name	Seed origin	Isolate	
			IRN TRa6	IRN SS5
Asteraceae				
Lactuca sativa cv. Salinas 88	Lettuce	Australia	(LI/M, D, S)[c]	(LI/M, D, S)
Balsaminaceae				
Impatiens balsamina	Garden balsam	Iran	NLL/BSS,PSF	NLL/BSS, PSF
Brassicaceae				
Brassica chinensis cv. Choyo	Qing geng cai	Denmark	LI/VC, M	LI/M
B. juncea cv. Hakarashina	Mustard	Italy	NLL/M, LD	NLL/VN, LD, S
B. napus cv. Otsubu	Oilseed rape	Japan	NLL/-	NLL/-
B. napus cv. PF	Oilseed rape	Germany	-/-	-/-
B. narinosa cv. Tatsuai	Rosette pakchoi	Australia	LI/M	LI/M, S
B. oleracea var. *botrytis* cv. Snow crown	Cauliflower	Australia	-/-	-/-
B. oleracea var. *botrytis* cv. Snow queen	Cauliflower	Chile	-/-	-/-
B. oleracea var. *botrytis* TN-84-17[a]	Cauliflower	Iran	-/-	-/-
B. oleracea var. *capitata*[b]	White cabbage	Iran	-/-	-/-
B. oleracea var. *capitata* cv. Ryozan-2 go	Green cabbage	Japan	-/-	-/-
B. oleracea var. *capitata* cv. Shinsei	Green cabbage	Japan	-/-	-/-
B. oleracea var. *capitata* cv. Yalova-1	White cabbage	Turkey	-/-	-/-
B. oleracea var. *capitata* cv. Zencibasi	Red cabbage	Turkey	-/-	-/-
B. oleracea var. *gongylodes*[b]	Kohlrabi	Iran	-/-	-/-
B. oleracea var. *italica* cv. Challenger	Broccoli	Japan	-/-	-/-
B. oleracea var. *italica* cv. cv. Endever	Broccoli	Japan	-/-	-/-
B. oleracea var. *italica* cv. Pixcel	Broccoli	Chile	-/-	-/-
B. pekinensis cv. Kyoto-3go	Chinese cabbage	Japan	CLL/-	NLL/-
B. pekinensis cv. Nozaki-1go	Chinese cabbage	Japan	-/-	-/-
B. rapa[b]	Turnip	Iran	LI/VC, SM, LD	LI/VC, SM
B. rapa cv. Ada-202	Turnip	Turkey	(LI/M)	(LI/M)
B. rapa cv. Hakatasuwari	Turnip	Japan	LI/M	LI/M
B. rapa TN-85-104[a]	Turnip	Iran	LI /VC, M	LI/VC, M
B. rapa TN-85-105[a]	Turnip	Iran	LI /VC, M	NLL/VC,M, VY
B. rapa TN-85-106[a]	Turnip	Iran	NLL/VC, M	LI/VC, M
B. rapa TN-85-107[a]	Turnip	Iran	NLL/VC, M	NLL/VC, M
B. rapa TN-85-119[a]	Turnip	Iran	NLL/VC, M	LI/VC, M
Eruca sativa cv. unknown	Rocket	Iran	(NLL/M)	(NLL/M)
E. sativa cv. Balikesir	Rocket	Turkey	(LI/M)	LI/M
E. sativa cv. Izmir	Rocket	Turkey	(LI/M)	(LI/M)
E. sativa cv. Odyssey	Rocket	Italy	(LI/M)	(LI/M)
Hirschfeldia incana	Shortpod mustard	Iran	NLL/VC, M	NLL/VC, M
Lepidium sativum	Garden cress	Iran	NLL/M	NLL/M, LD
Matthiola incana	Stock	Iran	LI/Mo	LI/Mo
Raphanus sativus[b]	Small radish	Iran	-/-	-/-
R. sativus cv. Akimasari	Japanese radish	Japan	-/-	-/-
R. sativus cv. Taibyo-sobutori	Japanese radish	New Zealand	-/-	-/-
R. sativus TN-91-33[a]	Small radish	Iran	-/-	-/-
R. sativus TN-91-49[a]	Small radish	Iran	-/-	-/-
Rapistrum rugosum	Annual bastard cabbage	Iran	LI/SM, LD	LI/M, Ru
Sinapis alba	White mustard	Iran	LI/M	LI/M
S. arvensis	Charlock mustard	Iran	NLL/M	LI/M
Sisymbrium irio	London rocket	Iran	LI/M	LI/M
S. loeselii	Small tumbleweed-mustard	Iran	(LI/M)	(LI/M, S)

Table 2 (continued)

Plant	Common name	Seed origin	Isolate IRN TRa6	IRN SS5
Chenopodiaceae				
Chenopodium quinoa	Quinoa	UK	CLL/CLL	CLL/CLL
Solanaceae				
Nicotiana benthamiana	Tobacco	Japan	LI/M	LI/M
N. glutinosa	Tobacco	Germany	(NL/NS,TN,D)	(LI/CS,S,TN,D)
Petunia hybrida	Petunia	Iran	LI/VY, CS	LI/M, Ru, VY, S

[a] Iranian local population with almost unique phenotypic properties.

[b] Iranian local seed obtained from Karaj (Tehran province).

[c] Reaction of inoculated leaves/upper leaves. At least three plants were inoculated.

BSS = Black stem streak, CLL = Chlorotic local lesion, CS = Chlorotic spots, D = Death, LD = Leaf deformation, LI = Latent infection, M = Mosaic, Mo = Mottle, NLL = Necrotic local lesion, PSF = Pod set failure, Ru = Rugose, S = Stunting, SCS = Small chlorotic spot, SL = Symptomless, SM = Severe mosaic, TN = Top necrosis, VC = Vein clearing, VY = Vein yellows, - = Not infected, (); Occasionally.

ALIGN (kindly supplied by Georg Weiller, Australian National University, Canberra, Australia) to maintain the degapped alignment of the encoded amino acids and the aligned sub-sequences were then reassembled to form complete sequences of 9321 nt. The aligned sequences were first checked for incongruent relationships that might have resulted from recombination, using RDP (Martin and Rybicki 2000), GENECONV (Sawyer 1999), BOOTSCAN (Salminen et al. 1995), MAXCHI (Maynard Smith 1992), CHIMAERA (Posada and Crandall 2001) and SISCAN programmes (Gibbs et al. 2000) in RDP3 (Martin et al. 2005). These analyses were done using default settings for the different detection programmes and a Bonferroni corrected *p*-value cut-off of 0.05 or 0.01. Next, all sequences that had been identified as likely recombinants, together with all those used in this study, were checked again using the original PHYLPRO (Weiller 1998) and SISCAN version 2, with all nucleotide sites, as well as with synonymous and non-synonymous sites separately. Both 100 and 50 nt slices of all sequences were checked for evidence of recombination using these programmes. These analyses also assessed which non-recombinant sequences had regions that were closest to regions of the recombinant sequences and hence indicated the likely lineages that provided those regions of the recombinant genomes. For simplicity, we called these the 'parental isolates' of recombinants, although in reality they were just the most closely related sequences in the set we were analysing.

Phylogenetic analyses

Possible genetic relationships and phylogenetic grouping of the sequences under study were investigated by two methods: the maximum-likelihood (ML) algorithm of TREEPUZZLE version 5.0 (Strimmer and von Haeseler 1996; Strimmer et al. 1997) and the neighbour-joining (NJ) algorithm of PHYLIP version 3.5 (Felsenstein 1993). For ML analyses, 1,000 puzzling steps were calculated using the Hasegawa-Kishino-Yano model of substitution (Hasegawa et al. 1985). For NJ analyses, distance matrices were calculated by DNADIST with the Kimura two-parameter option (Kimura 1980) and trees were constructed from these matrices by the NJ method (Saitou and Nei 1987). A bootstrap value for each internal node of the NJ trees was calculated using 1,000 random resamplings with SEQBOOT (Felsenstein 1985). The calculated trees were displayed using TREEVIEW (Page 1996). Multiple sequence alignment was performed using TRANSALIGN and CLUSTAL X, as above, to ensure that the alignment was congruent with the encoded amino acids. The separate gene regions were then reassembled to form sequences of 8,988 nts. Sequences of two JYMV and one ScMV isolates as well as the CP sequence of one NYSV isolate were used as outgroups to construct a phylogenetic tree of the concatenate regions. The nucleotide and amino acid similarities were estimated using the Kimura 2-parameter (Kimura 1980) and

Dayhoff PAM250 matrices (Dayhoff et al. 1983), respectively.

Results

Surveys and detection of TuMV by DAS-ELISA

A total of 532 samples collected from Brassicaceae weeds during 2003–2008 were tested by DAS-ELISA (Table 1). Among those tested, 340 samples (64%) were found to be infected with TuMV. Infection rates of this virus varied either among collection regions or Brassicaceae weed species. The highest infection rate of TuMV was found for *R. rugosum* (82 infected plants/82 collected plants, 100%), followed by *S. loeselii* (235/322, 73%), *S. irio* (22/76, 28.9%) and *H. incana* (1/5, 20%), whereas no infection was detected in weeds of *B. deflexa* and *S. arvensis*. Occurrence of the virus in weeds was found in most crop and ornamental brassica-growing areas of Iran, including Hamedan, Kerman, Khorasan, Markazi, Semnan and Tehran provinces (Table 1, Fig. 1).

Host range and symptoms

The 19 isolates were chosen among 340 TuMV-infected samples considering origin of location and host plant species. The isolates from *R. rugosum*, *S. loeselii* and *S. irio* were examined to determine their specific host range; four came from Hamedan, three were from Kerman, three were from Khorasan, one was from Markazi, three were from Semnan and five were from Tehran (Table 1). All the 19 isolates infected many *Brassica* plants systemically but did not infect *R. sativus*; these belong to the B-infecting host type (data not shown). Because host reactions of the isolates studied were very similar and the highest infection rates were found for *R. rugosum* and *S. loeselii* plants, two isolates IRN TRa6 and IRN SS5 in Tehran and Semnan provinces were selected for a more detailed study of host range using more variable test plants (Table 2).

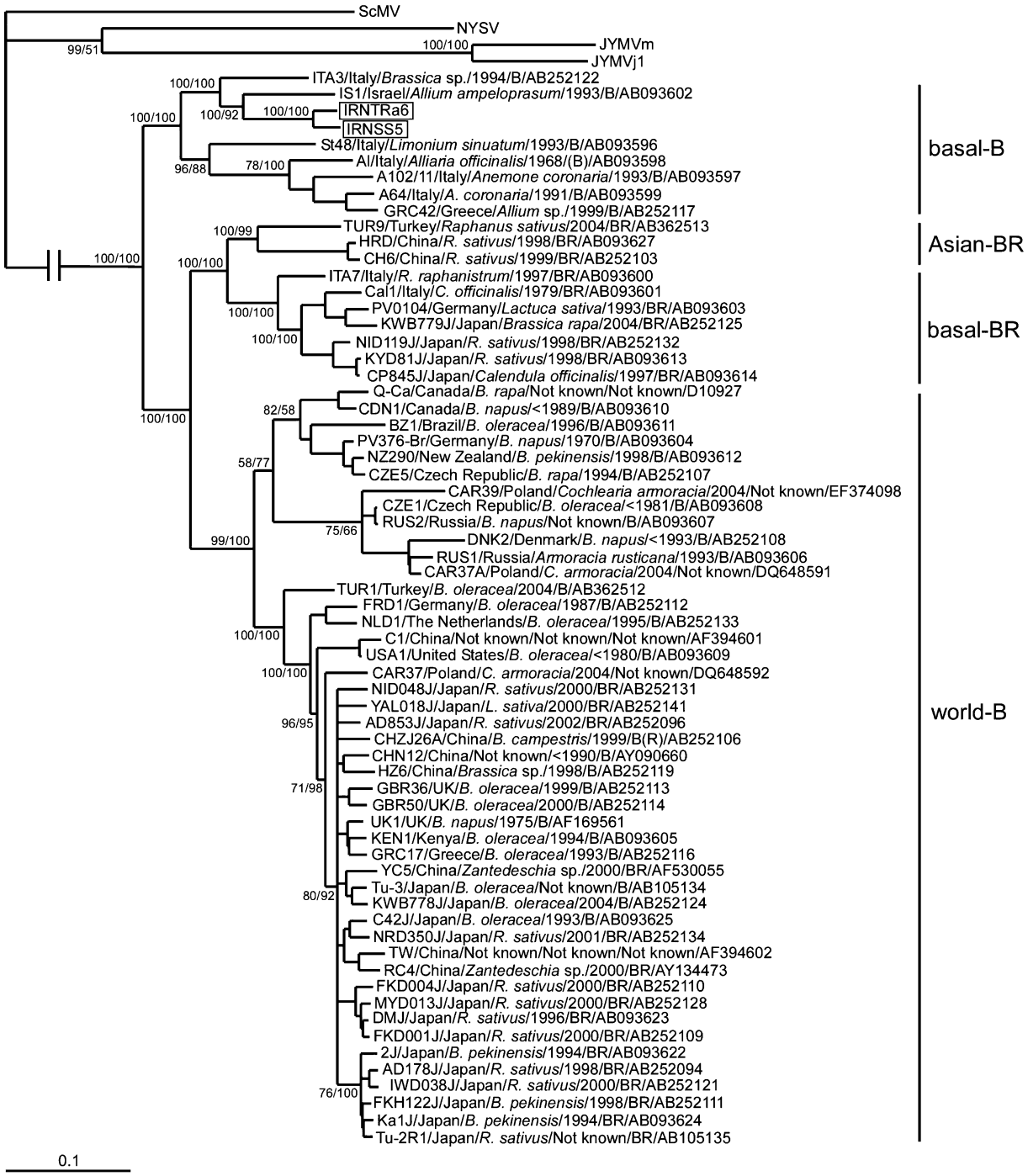
Many *Brassica* plant species were used for host range examination. The two isolates infected *B. chinensis* (Qing geng cai), *B. juncea* (mustard) and *B. rapa* (turnip) but not *B. oleracea* (broccoli, cauliflower and cabbage). IRN TRa6 produced necrotic local lesions on the inoculated leaves and did not systemically infect *B. pekinensis* whereas IRN SS6 did not infect this plant. On the other hand, IRN SS6 produced necrotic local lesions on the inoculated leaves of *B. napus* cv. Otsubu whereas IRN TRa6 did not. Both isolates infected *E. sativa* (rocket), *R. rugosum* (annual bastard cabbage) and *S. loeselii* (small tumbleweed-mustard) with similar symptoms. Thus, despite minor differences in pathogenicity for some of the tested Brassicaceae and other hosts, the two isolates showed similar biological characterisation and are categorised in the B-infecting host type.

Nucleotide sequences

The genomic nucleotide sequence of IRN TRa6 and IRN SS5 isolates, excluding the 5′ end primer sequences of 26 nt, was determined. The sequences of both isolates were 9807 nt long. The results indicated that these sequences encompass regions encoding P1, HC-Pro, third protein (P3), 6Kda 1 protein (6K1), cylindrical inclusion protein (CI), 6Kda 2 protein (6K2), genome linked viral protein (VPg), nuclear inclusion proteinase protein (NIa-Pro), NIb and CP genes with 1086, 1374, 1065, 156, 1932, 159, 576, 729, 1551 and 864 nts, respectively. All of the motifs reported for different potyvirus genes and encoded proteins were found. The genomic sequences of IRN TRa6 and IRN SS5 determined in this study were deposited in DDBJ/EMBL/GenBank databases (the accession numbers AB440238 and AB440239).

Identification of recombinants

The polyprotein sequences were assessed for evidence of recombination. The 5′ and 3′ non-coding regions were omitted because polyprotein sequences needed to be examined using not only all of the nucleotide sites but also the synonymous and non-synonymous sites separately. After all gaps and nucleotides homologous to them had been removed from the aligned sequences, the likely recombination sites were assessed using RDP3. Each of the identified sites was examined individually and a phylogenetic approach was used to verify the parent/donor assignments made by RDP3. Having examined all sites with an associated p value of $<1\times10^{-6}$ (i.e., the most obvious events), the intralineage recombinants (parents from the same major lineage) were retained and the interlineage recombinants (parents from different



major lineages) removed by treating the identified recombination sites as missing data in subsequent analyses. Moreover, the 'phylogenetic profiles' of the polyprotein sequences, which were examined by PHYLPRO and SISCAN version 2 programmes, were then used to check for evidence of recombination, not only in the total nucleotide sites but also in synonymous and non-synonymous sites separately. This complex approach was adopted to find all of the recombination sites as well as to decrease the possibility of obtaining false evidence of recombination. Finally, no 'clear' recombination (p value of $<1\times10^{-6}$) was found in the

◀ **Fig. 2** A maximum likelihood (ML) phylogenetic tree showing the relationship among weed isolates of *Turnip mosaic virus* (TuMV) with the other isolates. The tree was constructed from the polyprotein sequences of all isolates, excluding the interlineage recombinants identified in this study and in earlier studies (Ohshima et al. 2007; Korkmaz et al. 2008). Numbers at each node indicate the percentage of supporting puzzling steps (or bootstrap samples) in ML and NJ methods, respectively (only values >50 are shown). Horizontal branch lengths are drawn to scale with the bar indicating 0.1 nt replacements per site. The homologous sequences of two isolates (mild and j1) of *Japanese yam mosaic virus* (JYMV), an isolate of *Scallion mosaic virus* (ScMV) and an isolate of *Narcissus yellow stripe virus* (NYSV) were used as the outgroup. For details of the phylogenetic groups, basal-B, basal-BR, Asian-BR, and world-B, see Ohshima et al. (2002). The name of each isolate, its country of origin, original host plant, year of isolation, host type and accession code in the international gene sequence data base are listed

genome of IRN SS5 and IRN TRa6 and the isolates seemed not to be recombinants. On the other hand, the full genomic nucleotide sequences of two Polish isolates of CAR37 (DQ648592) and CAR39 (EF374098), recently deposited in the international gene sequence databases (Kozubek et al. 2007), were found to have recombination sites in P1 and P3 genes, respectively; a recombination site in the middle of P1 gene of CAR37 isolate was detected by *p* values using the RDP (5.7×10^{-26}), GENECONV (2.2×10^{-10}), BOOTSCAN (1.7×10^{-25}), MAXCHI (1.6×10^{-7}), CHIMAERA (3.2×10^{-8}), and SISCAN (5.5×10^{-14}) programmes, and a recombination site in the N terminal region of P3 gene of CAR39 isolate was detected by *p* values using the RDP (3.2×10^{-132}), GENECONV (1.5×10^{-126}), BOOTSCAN (1.7×10^{-135}), MAXCHI (8.0×10^{-40}), CHIMAERA (1.3×10^{-27}), and SISCAN (8.6×10^{-72}) programmes of the RDP3 software. These two sites were 'clear' intralineage (parental sequences from identical major genogroup) recombination sites of world-B parents. Other recombination sites found in the present study were the same as those found in our earlier studies (Ohshima et al. 2002, 2007; Korkmaz et al. 2008). Altogether, 71 out of 99 sequences were found to be recombinants.

Phylogenetic relationships

Trees were initially calculated from the genome sequences of the 99 isolates, including all of the recombinants identified in this study. However, there were inconsistencies in and poor bootstrap support for some lineages in the resulting trees, as found previously (Ohshima et al. 2002, 2007). Therefore, the trees were recalculated from the genomes of only 65 isolates, discarding the interlineage (parental sequences from different major genogroups) recombinants. The relationships of these isolates were investigated by ML and NJ methods and the ML tree is shown in Fig. 2. All of the trees partitioned most of the sequences into the same four consistent groups: basal-B, basal-BR, Asian-BR, and world-B, as reported previously (Ohshima et al. 2007). The basal-B group was a sister group to all others in the ML and NJ phylogenies and was supported by high bootstrap values. IRN TRa6 and IRN SS5 isolates from *R. rugosum* and *S. loeselii*, respectively, fell into the basal-B group, confined to B-infecting host type isolates from Mediterranean regions of Greece, Israel and Italy. As both isolates were found to be non-

Table 3 Comparison of nucleotide and amino acid similarities between Iranian isolates of *Turnip mosaic virus* with the other exotic isolates belonging to basal-B group[a]

Isolate	IRN TRa6	IRN SS5	ITA3	IS1	St48	Al	A102/11	A64	GRC42
IRN TRa6	–	0.0404	0.1423	0.1384	0.1760	0.1876	0.1828	0.1849	0.1902
IRN SS5	0.0110	–	0.1430	0.1410	0.1735	0.1882	0.1851	0.1867	0.1907
ITA3	0.0395	0.0422	–	0.1429	0.1772	0.1893	0.1841	0.1922	0.1925
IS1	0.0384	0.0428	0.0466	–	0.1797	0.1808	0.1751	0.1840	0.1873
St48	0.0601	0.0591	0.0610	0.0674	–	0.1784	0.1753	0.1762	0.1782
Al	0.0668	0.0665	0.0676	0.0696	0.0627	–	0.1087	0.1036	0.1048
A102/11	0.0624	0.0635	0.0593	0.0616	0.0567	0.0411	–	0.0895	0.0925
A64	0.0691	0.0706	0.0707	0.0684	0.0665	0.0297	0.0395	–	0.0471
GRC42	0.0671	0.0686	0.0673	0.0675	0.0635	0.0283	0.0379	0.0094	–

[a] The nucleotide (below diagonal) and amino acid (above diagonal) similarities were assessed using Kimura two-parameter (Kimura 1980) and Dayhoff PAM250 matrices (Dayhoff et al. 1983). Isolates ITA3, St48, Al, A102/11 and A64 were collected in Italy, isolates IS1 and GRC42 were collected in Israel and Greece, respectively.

recombinant, we calculated the nucleotide and amino acid similarities between the other two isolates in the basal-B group using Kimura 2-parameter (Kimura 1980) and Dayhoff PAM250 matrix (Dayhoff et al. 1983) (Table 3). The similarities showed that the IRN TRa6 and IRN SS5 isolates were most closely related to each other, and distinct from the other isolates from Mediterranean regions in the basal-B group.

Discussion

In this study, the Brassicaceae reservoir weed hosts of TuMV were surveyed during the period 2003–2008 and the virus was found to be widespread in mid-Eurasia, Iran. However, the incidence of TuMV differed among plant species surveyed, and *R. rugosum* and *S. loeselii*, which are reported for the first time as natural weed hosts, were found to be the most dominant weed hosts of TuMV. These weeds were found to play an important role in virus ecology and epidemiology in mid-Eurasia.

Preliminary results of host-range studies revealed that all the 19 weed isolates belong to the B-infecting host type. The detailed host reactions of IRN TRa6 and IRN SS5 isolates showed that they are similar in their biological characteristics on a number of plant species and cultivars tested (Table 2). However, these isolates did not infect *B. oleracea* varieties including cabbage, cauliflower and broccoli but did infect some other *Brassica* plants such as turnip and mustard, suggesting that the viruses from *R. rugosum* and *S. loeselii* may be the source of infection for the latter cultivated *Brassica* plants in Iran.

The simplest interpretation of the phylogenetic analyses and the genetic structure of the populations between worldwide isolates in earlier studies (Ohshima et al. 2002; Tomimura et al. 2004) is that the ancestral TuMV population was the basal-B group and the most diverse, consists of isolates only from Eurasia and was a sister group to all the others, and would be from southeast European, Asia Minor or mid-Eurasian regions. Both IRN TRa5 and IRN SS5 isolates collected in mid-Eurasia fell into basal-B group (Fig. 2); although evolutionary analyses using a large number of the isolates from Iran are necessary, our results may support the conclusion.

Recombination plays a major role in plant RNA and DNA virus variability and evolution. In the case of potyviruses, a total of 32 out of 51 (63%) of the full *Potato virus Y* genomic sequences (Ogawa et al. 2008), and a total of 68 out of 92 (74%) of the full TuMV genomic sequences were found to be recombinants (Ohshima et al. 2007). In this study, 71 out of 99 sequences were found to be recombinants (data not shown), and neither of two Iranian basal-B isolates seemed to be recombinants. Our results mirror our earlier studies of TuMV that recombinants from parental Asian-BR and world-B sequences in East Asia were relatively easy to find, whereas it was more difficult to find recombinants from ancestral basal-B sequences (Tomimura et al. 2003; Ohshima et al. 2007). As we have little information of the population for basal-B group, the similarities of nucleotide and amino acid sequences between two Iranian and other basal-B isolates were calculated (Table 2). The result showed that the two Iranian isolates are distinct from the basal-B isolates from Italy, Israel and Greece. Therefore, Iranian and other Mediterranean isolates seemed to be genetically distinct, as shown in the ML tree in Fig. 2.

Although Iranian isolates belonging to basal-B group were found in the mid-Eurasia population, we do not know yet whether these are dominant isolates in this region; evolutionary comparisons of a large number of isolates from southeast Europe, Asia Minor and mid-Eurasia with representative worldwide isolates are necessary to determine this. However, to our knowledge, the present study shows for the first time the wide distribution and evolutionary relationships of TuMV from weeds collected in the mid-Eurasian region of Iran.

Acknowledgements We are grateful to S. Sajedi, Botanic Department of Plant Pests and Diseases Research Institute (PPDRI), Iran, for identification of weed species, to Dr. S. Korkmaz, Department of Plant Protection, Faculty of Agriculture, University of Canakkale Onsekiz Mart, Turkey, for preparing some seeds used in the study, and to Prof. U. Melcher, Department of Biochemistry and Molecular Biology, Oklahoma State University, USA, for his kindly reading the manuscript. This work was supported by a grant from PPDRI of Iran and Grant-in-Aid for Scientific Research (B) No. 18405022 from the Japan Society for the Promotion of Science.

References

Clark, M. F., & Adams, A. M. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for detection of plant viruses. *The Journal of General Virology*, *34*, 475–483. doi:10.1099/0022-1317-34-3-475.

Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983). Establishing homologies in protein sequences. *Methods in Enzymology*, *91*, 524–545. doi:10.1016/S0076-6879(83)91049-2.

Farzadfar, S., Ohshima, K., Pourrahim, R., Golnaraghi, A. R., Jalali, S., & Ahoonmanesh, A. (2005). Occurrence of *Turnip mosaic virus* on ornamental crops in Iran. *Plant Pathology*, *54*, 261. doi:10.1111/j.1365-3059.2004.01148.x.

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (2005). *Virus taxonomy: Classification and nomenclature of viruses, eighth report of the international committee on taxonomy of viruses*. San Diego: Elsevier Academic.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution; International Journal of Organic Evolution*, *39*, 783–791. doi:10.2307/2408678.

Felsenstein, J. (1993). PHYLIP (Phylogeny interference package), Version 3.5. Department of Genetics, University of Washington, Seattle.

Gibbs, M. J., Armstrong, J. S., & Gibbs, A. J. (2000). Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics (Oxford, England)*, *16*, 573–582. http://www.anu.edu.au/BoZo/software/. doi:10.1093/bioinformatics/16.7.573

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*, 160–174. doi:10.1007/BF02101694.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., & Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences*, *23*, 403–405. doi:10.1016/S0968-0004(98)01285-7.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*, 111–120. doi:10.1007/BF01731581.

Korkmaz, S., Tomitaka, Y., Onder, S., & Ohshima, K. (2008). Occurrence and molecular characterization of Turkish isolates of *Turnip mosaic virus*. *Plant Pathology*. doi:10.1111/j.1365-3059.2008.01902.x.

Kozubek, E., Irzykowski, W., & Lehmann, P. (2007). Genetic and molecular variability of a *Turnip mosaic virus* population from horseradish (*Cochlearia armoracia* L.). *Journal of Applied Genetics*, *48*(3), 295–306.

Martin, D., & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics (Oxford, England)*, *16*, 562–563. doi:10.1093/bioinformatics/16.6.562.

Martin, D., Williamson, C., & Posada, D. (2005). RDP2: recombination detection and analysis from sequence alignment. *Bioinformatics (Oxford, England)*, *21*, 260–262. doi:10.1093/bioinformatics/bth490.

Maynard Smith, J. (1992). Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, *34*, 126–129.

Ogawa, T., Tomitaka, Y., Nakagawa, A., & Ohshima, K. (2008). Genetic structure of a population of *Potato virus Y* inducing potato tuber necrotic ringspot disease in Japan; comparison with North American and European populations. *Virus Research*, *131*, 199–212. doi:10.1016/j.virusres.2007.09.010.

Ohshima, K., Tomitaka, Y., Wood, J. T., Minematsu, Y., Kajiyama, H., Tomimura, K., et al. (2007). Patterns of recombination in *Turnip mosaic virus* genomic sequences indicate hotspots of recombination. *The Journal of General Virology*, *88*, 298–315. doi:10.1099/vir.0.82335-0.

Ohshima, K., Yamaguchi, Y., Hirota, R., Hamamoto, T., Tomimura, K., Tan, Z., et al. (2002). Molecular evolution of *Turnip mosaic virus*; evidence of host adaptation, genetic recombination and geographical spread. *The Journal of General Virology*, *83*, 1511–1521.

Page, R. D. M. (1996). TreeView: an application to display phylogenetic trees on personal computer. *Computer Applications in the Biosciences*, *12*, 357–358.

Posada, D., & Crandall, K. A. (2001). Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 13757–13762. doi:10.1073/pnas.241370698.

Provvidenti, R. (1996). *Turnip mosaic potyvirus*. In A. A. Brunt, K. Crabtree, M. J. Dallwitz, A. J. Gibbs, & L. Watson (Eds.), *Viruses of plants* (pp. 1340–1343). Wallingford: CAB International.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*, 406–425.

Salminen, M. O., Carr, J. K., Burke, D. S., & McCutchan, F. E. (1995). Identification of breakpoints in intergenotypic recombinants of HIV type 1 by Bootscanning. *AIDS Research and Human Retroviruses*, *11*, 1423–1425.

Sawyer, S. A. (1999). GENECONV: A computer package for the statistical detection of gene conversion. Distributed by the author. Department of Mathematics, Washington University in St. Louis, available at http://www.math.wustl.edu/~sawyer.

Strimmer, K., & von Haeseler, A. (1996). Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution*, *13*, 964–969.

Strimmer, K., Goldman, N., & von Haeseler, A. (1997). Bayesian probabilities and quartet puzzling. *Molecular Biology and Evolution*, *14*, 210–211.

Tomimura, K., Gibbs, A. J., Jenner, C. E., Walsh, J. A., & Ohshima, K. (2003). The phylogeny of *Turnip mosaic virus*; comparisons of thirty-eight genomic sequences reveal a Eurasian origin and a recent 'emergence' in east Asia. *Molecular Ecology*, *12*, 2099–2111. doi:10.1046/j.1365-294X.2003.01881.x.

Tomimura, K., Spak, J., Katis, N., Jenner, C. E., Walsh, J. A., Gibbs, A. J., et al. (2004). Comparisons of the genetic structure of populations of *Turnip mosaic virus* in west and east Eurasia. *Virology*, *330*, 408–423. doi:10.1016/j.virol.2004.09.040.

Tomlinson, J. A. (1987). Epidemiology and control of virus diseases of vegetables. *The Annals of Applied Biology*, *110*, 661–681. doi:10.1111/j.1744-7348.1987.tb04187.x.

Urcuqui-Inchima, S., Haenni, A-L., & Bernardi, F. (2001). Potyvirus proteins: a wealth of functions. *Virus Research*, *74*, 157–175. doi:10.1016/S0168-1702(01)00220-9.

Walsh, J. A., & Jenner, C. E. (2002). *Turnip mosaic virus* and the quest for durable resistance. *Molecular Plant Pathology*, *3*, 289–300. doi:10.1046/j.1364-3703.2002.00132.x.

Weiller, G. F. (1998). Phylogenetic profiles: a graphical method for detecting genetic recombinations in homologous sequences. *Molecular Biology and Evolution*, *15*, 326–335.